

Averaging and sampling for magnetic-observatory hourly data

J. J. Love¹, V. C. Tsai², and J. L. Gannon¹

¹Geomagnetism Program, US Geological Survey, Denver, CO, USA

²Earthquake Hazards Program, US Geological Survey, Denver, CO, USA

Received: 21 May 2010 – Revised: 25 October 2010 – Accepted: 5 November 2010 – Published: 12 November 2010

Abstract. A time and frequency-domain analysis is made of the effects of averaging and sampling methods used for constructing magnetic-observatory hourly data values. Using 1-min data as a proxy for continuous, geomagnetic variation, we construct synthetic hourly values of two standard types: instantaneous “spot” measurements and simple 1-h “boxcar” averages. We compare these average-sample types with others: 2-h average, Gaussian, and “brick-wall” low-frequency-pass. Hourly spot measurements provide a statistically unbiased representation of the amplitude range of geomagnetic-field variation, but as a representation of continuous field variation over time, they are significantly affected by aliasing, especially at high latitudes. The 1-h, 2-h, and Gaussian average-samples are affected by a combination of amplitude distortion and aliasing. Brick-wall values are not affected by either amplitude distortion or aliasing, but constructing them is, in an operational setting, relatively more difficult than it is for other average-sample types. It is noteworthy that 1-h average-samples, the present standard for observatory hourly data, have properties similar to Gaussian average-samples that have been optimized for a minimum residual sum of amplitude distortion and aliasing. For 1-h average-samples from medium and low-latitude observatories, the average of the combination of amplitude distortion and aliasing is less than the 5.0 nT accuracy standard established by Intermagnet for modern 1-min data. For medium and low-latitude observatories, average differences between monthly means constructed from 1-min data and monthly means constructed from any of the hourly average-sample types considered here are less than the 1.0 nT resolution of standard databases. We recommend that observatories and World Data Centers continue the standard practice of reporting simple 1-h-average hourly values.

Keywords. Electromagnetics (Measurement and standards)

1 Introduction

Since magnetic observatories first began operating around the world in the middle of the 19th century, one of their most important products has been hourly data values. Historical hourly observatory data are useful for a variety of applications (e.g. Parkinson, 1983; Courtillot and Le Mouél, 1988; Barraclough et al., 1992; Prölss, 2004), including studying geomagnetic secular variation originating in the core (e.g. Sabaka et al., 2004), exploring the electrical conductivity of the mantle (e.g. Egbert et al., 1992), mapping electric currents in the ionosphere (e.g. Campbell, 1989), measuring the intensity of magnetospheric storms (e.g. Karinen and Mursula, 2005), and estimating long-term solar-terrestrial interaction (e.g. Macmillan and Droujinina, 2007). In conducting these, and other types of analyses, it is important, and sometimes even essential, to have an understanding of the effects of amplitude distortion and aliasing caused by averaging and sampling procedures used in the production of hourly data.

The means of acquiring hourly observatory data have evolved with the advancement of measurement technology and in response to researcher demands for data that meet increasingly stringent quality standards. The oldest observatory hourly data were obtained by on-site personnel using instruments that permitted essentially instantaneous visual measurement of magnetic-field direction and intensity. With the introduction of photographic recording systems (Brooke, 1847), the process of acquiring hourly data became semi-automated. Following a daily schedule, an observatory worker would develop the photographic paper, and, then, using an etched piece of glass, directly measure the amplitudes of the time series traces recording magnetic-field variation. At first these were “spot” measurements, similar, in some respects, to direct instantaneous visual measurements. In the early 20th century, following the recommendation of Schmidt (1908), observatories began to report hourly means, with 1-h average values recorded once per hour (Chapman and Bartels, 1962, Ch. 2.14). These were either directly estimated from the paper photographic records,



Correspondence to: J. J. Love
(jlove@usgs.gov)

or, more formally, they were calculated from multiple sub-hour spot measurements. The switch from making spot measurements to producing average values came gradually: in Great Britain (Eskdalemuir) it was made in 1912 (Meteorological Office, 1914, p. 68), in the United States at all observatories in 1915 (e.g. Hazard, 1918, p. 5), and at some observatories in France (Chambon la Forêt) not until 1972 (Fouassier and Chulliat, 2009, p. 87). Today, constructing hourly means is done by computer, averaging digital 1-min data acquired by electronic systems.

Because the historical hourly observatory data are a mixture of spot measurements and hourly averages, their properties have changed over time, and their time series can be described as being “inhomogeneous”. Spot sampling can sometimes, and fortuitously, give a good record of short-duration transient magnetic-field variation. Spot sampling can also sometimes, and unfortunately, miss transient variation altogether. It all depends on when and how frequently samples are taken relative to when and over what duration the field variation occurs. If we assume, as is reasonable, that spot measurements are collected independently of field variation, then numerous spot measurements collected over a long period of time will represent an unbiased statistical sampling of the amplitude range of magnetic-field variation. On the other hand, averaging reduces the amplitude of high-frequency variation, giving a time series that is smoother than actual magnetic-field variation. The statistical variance of discrete 1-h-average values will be less than that of the magnetic field’s actual variation, and it will be less than that of instantaneous, spot measurements. Since magnetic-field variation occurs over an extremely broad range of frequencies, discrete hourly sampling, be it done with spot measurement or with an averaging window, will result in aliasing, with high-frequency amplitudes being mapped into estimates of low-frequency amplitudes.

Researchers often use a combination of spot and hourly-average observatory data, as if, together, they represent a reliable long-term record of magnetic-field variation. It is, therefore, natural to ask: can we measure differences in the properties of hourly spot and hourly average values so that informed decisions can be made about using them? How do spot and hourly averages compare with other hypothetical average-sample types that might be proposed in the future as candidate observatory products? In seeking answers to these and other related questions, we examine the properties of hourly observatory data, measuring their variance, relative proportionality, correlation, autocorrelation, spectral power, amplitude distortion, and aliasing. Assuming that 1-min data are a good representation of continuous magnetic-field variation, we construct several types of synthetic hourly average-samples from 1-min data collected from observatories situated at different latitudes. We employ standard methods of continuous (e.g. Bracewell, 1978; Kanasewich, 1981) and statistical (e.g. Lee, 1960; Bendat and Piersol, 2000) time-series analysis.

2 Data

In Table 1 we summarize the data we use. The 1-min data are definitive (processed and calibrated) magnetic-vector values collected at observatories (Jankowski and Sucksdorff, 1996; Love, 2008) that meet Intermagnet standards (Keridge, 2001; Rasson, 2007); we obtained the 1-min data from the Intermagnet website (www.intermagnet.org) for observatories situated at high (Barrow BRW), medium (Chambon la Forêt CLF), and low (Huancayo HUA) geomagnetic latitudes. They record a variety of magnetic-field variation, with high (low) latitude variation dominated by active auroral (equatorial) electrojets (e.g. Parkinson, 1983; Prölss, 2004), and medium-latitude variation being relatively more quiescent. Each 1-min magnetic-vector datum was formed from sub-minute electronic signals acquired from a tri-axial fluxgate magnetometer. The signals were both analog and digitally filtered, effectively eliminating aliasing from variation with periods of less than 1-min (frequencies greater 0.0167 Hz). Additional measurements, made once a week or so from a reference pier at each observatory site, were combined with the fluxgate measurements to construct time series that have an accuracy, measured in absolute terms over many years, of better than 5.0 nT (0.50′). Each 1-min datum has a time stamp that is centered on the top of the universal-time (UT) minute (HR:MN:SC, 00:00:00, 00:01:00, etc.). For BRW and HUA, we use almost a complete solar cycle of data (1998.0–2009.0), and for CLF we use almost two solar cycles of data (1991.0–2009.0).

The hourly values we use were acquired at Chambon la Forêt (CLF) for the years 1954.0–1990.0 (more than three solar cycles) and at Eskdalemuir (ESK) for the years 1917.0–1947.0 (almost three solar cycles). These are available in computer-readable format from the website of the World Data Center “CE” in Copenhagen (now Edinburgh), having been either transcribed from printed observatory year-book values or constructed from digital minute data. Prior to 1972.0 (1932.0) the CLF (ESK) data are spot samples (1-h means) centered on the top of the Greenwich hour (00:00:00, 01:00:00, etc.) (Fouassier and Chulliat, 2009; Meteorological Office, 1933). After 1972.0 (1932.0) they are, for both CLF and ESK, 1-h means centered on the bottom of each hour (00:30:00, 01:30:00, etc.) (Fouassier and Chulliat, 2009; Meteorological Office, 1934), as is now the standard among other observatories. Martini and Mursula (2006) and Svalgaard and Cliver (2007) have noted that the WDC-CE pre-1932 ESK holdings are an interpolation of hourly values centered at the top of each hour (00:00:00, 01:00:00, etc.) – adjacent values in time have been averaged together to give 2-h average-samples, one per hour and centered on the bottom of the hour¹. We use the CLF-CE data to investigate

¹This was probably done with the intent of making the data easier to use, but the result is actually the opposite for anyone wishing to conduct a detailed analysis. The pre-1932.0 ESK observatory

Table 1. Summary of observatories and data used. Geomagnetic latitudes are calculated for a 2005 dipole.

Code	Observatory name	Region	Data used	Years	Geomag. Lat.	Present supporting institute
BRW	Barrow	Alaska	min Intermagnet	1998.0–2009.0	69.6°	US Geological Survey
CLF	Chambon la Forêt	France	min Intermagnet	1991.0–2009.0	49.8°	Institut de Physique du Globe de Paris
HUA	Huancayo	Perú	min Intermagnet	1998.0–2009.0	−1.8°	Instituto Geofísico
CLF	Chambon la Forêt	France	h WDC-CE	1954.0–1990.0	49.8°	Institut de Physique du Globe de Paris
ESK	Eskdalemuir	Britain	h WDC-CE	1917.0–1947.0	57.8°	British Geological Survey

differences between hourly spot samples and 1-h averages. We use the ESK-CE data to investigate the effects of heavy averaging.

For both the 1-min data and hourly values, we use geographic-polar (horizontal H , declination D , vertical Z) magnetic-vector components. When these are the components in the WDC database, then we use them without any modification, otherwise, if and when cartesian (north X , east Y) horizontal components are reported, then we calculate (D, H) appropriately. Data spikes are identified with a simple computer algorithm; these are removed and, along with occasional data gaps, filled with interpolated values. The British Geological Survey's annual-means database (www.geomag.bgs.ac.uk) contains a list of small step-offsets in the historical time series; some are also documented in observatory yearbooks. Most step offsets are obvious from inspection of the data; they are the result of moving the observatory's reference pier, changes in measurement method, or uncontrolled contamination. We insert complementary step-offsets into the historical observatory time series in order to bring them into continuity. The details of these adjustments do not significantly affect the results that follow.

3 Theory

The standard averaging and sampling methods that are used for constructing observatory hourly values are mathematically equivalent to passing a continuous magnetic-field time series through a zero-phase-delay linear filter, and, then, discretely sampling the output once per hour. In this section of mathematical review, we examine the properties of different averaging and sampling types in the dual domains of time and frequency. The average-samples are simple, and they include the standard spot and 1-h averages used for constructing hourly values, as well as other average-sample types useful for comparison.

yearbooks record data in XYZ coordinates, but the pre-1932.0 WDC-CE holdings are in HDZ coordinates – another indication that the data were manipulated at some point after first reporting. Furthermore, the pre-1932.0 ESK digital holdings of WDC Kyoto are different from those of WDC-CE. Indeed, the WDC-K data appear to be close to the original yearbook data, but there are some formatting problems with the WDC-format version of the ESK-K data.

For notation, we represent the time series of an arbitrary magnetic-vector component as $B(t)$, which is a function of time t , and which can stand for any of $H(t)$, $D(t)$, $Z(t)$, etc. We assume that $B(t)$ is well behaved: it is finite, continuous, and integrable. These qualities pertain to any natural magnetic-field time series measured over a finite duration of time and that might be generated from classical physical processes. Therefore, once the data have been detrended, a Fourier-type analysis can be reasonably pursued. We use a Fourier transformation \mathcal{F} having a “unitary” normalization for ordinary frequencies f , where $t = 1/f$. For a signal $B(t)$ in the time domain, its dual in the frequency domain, $\tilde{B}(f)$, is given by

$$\tilde{B}(f) = \mathcal{F}\{B(t)\} = \int_{-\infty}^{+\infty} B(t)e^{-2\pi ift} dt. \quad (1)$$

Inverse Fourier transformation is given by

$$B(t) = \mathcal{F}^{-1}\{\tilde{B}(f)\} = \int_{-\infty}^{+\infty} \tilde{B}(f)e^{2\pi ift} df. \quad (2)$$

The unitary normalization is used in the classic textbook by Bracewell (1978, p. 6); it is different from the non-unitary normalization for angular frequencies used by Lee (1960, p. 33). In most respects, this is a technical distinction, one that does not affect a qualitative understanding of the theory and results that follow.

3.1 Simple running averaging

We begin by considering the straightforward running average of a geomagnetic time series. With a rectangle-in-time or “boxcar” function, simple averaging is defined over a duration of time t_a by convolution of

$$\text{rect}(t; t_a) \equiv \begin{cases} 1 & |t| < t_a/2 \\ 0 & |t| > t_a/2 \end{cases} \quad (3)$$

with the natural, continuous magnetic-field time series $B(t)$,

$$\begin{aligned} B_a(t) &= \frac{1}{t_a} B(t) * \text{rect}(t; t_a) \\ &= \frac{1}{t_a} \int_{-\infty}^{+\infty} B(\zeta) \text{rect}(t - \zeta; t_a) d\zeta \\ &= \frac{1}{t_a} \int_{t-t_a/2}^{t+t_a/2} B(\zeta) d\zeta. \end{aligned} \quad (4)$$

Fourier transformation of the rectangle function in the time domain is a sinc function in the frequency domain,

$$\frac{1}{t_a} \mathcal{F}\{\text{rect}(t; t_a)\} = \frac{\sin(\pi f/f_a)}{\pi f/f_a} \equiv \text{sinc}(f; f_a) \quad (5)$$

(Bracewell, 1978, p. 389), where $f_a = 1/t_a$.

By the convolution theorem (Lee, 1960, p. 28; Bracewell, 1978, Ch. 3), the Fourier transform of convolution in the time (frequency) domain is equal to multiplication in the frequency (time) domain,

$$\mathcal{F}\{B(t) * \text{rect}(t; t_a)\} = \mathcal{F}\{B(t)\} \mathcal{F}\{\text{rect}(t; t_a)\}. \quad (6)$$

With this,

$$\mathcal{F}\{B_a(t)\} = \tilde{B}_a(f) = \tilde{B}(f) \text{sinc}(f; f_a), \quad (7)$$

where $\tilde{B}(f)$ is the frequency-domain dual of $B(t)$. The function $\text{sinc}(f; f_a)$ gives the mapping from $\tilde{B}(f)$ to $\tilde{B}_a(f)$; it is a “frequency-response” function (Bracewell, 1978, p. 179; Kanasewich, 1982, Ch. 5.3) that describes the amplitude distortion of the time series in the frequency domain due to averaging. It is useful, now, to define the amplitude differences

$$A_a(t) = B(t) - B_a(t) \quad \text{and} \quad \tilde{A}_a(f) = \tilde{B}(f) - \tilde{B}_a(f), \quad (8)$$

which measure the signal in the original, continuous time series that is not modelled by the running average. For a geomagnetic reference to related issues, see Chapman and Bartels (1962, Ch. 16.17).

3.2 Low-pass brick-wall filtering

A low-pass “brick-wall” filter is most easily defined in the frequency domain (Bracewell, 1978, p. 52; Kanasewich, 1981, Ch. 15): it has a flat response below some chosen truncation frequency $f_b/2$, and it admits nothing above that frequency. Because the brick-wall filter has these properties, it is sometimes described as “ideal”. It is prescribed by multiplication of the Fourier amplitudes by a rectangle-in-frequency function

$$\tilde{B}_b(f) = \tilde{B}(f) \text{rect}(f; f_b), \quad (9)$$

and, correspondingly, by convolution in the time domain with a sinc function,

$$\mathcal{F}^{-1}\{\tilde{B}_b(f)\} = B_b(t) = \frac{1}{t_b} B(t) * \text{sinc}(t; t_b). \quad (10)$$

The difference between the original, continuous time series and the filtered time series is equal to unresolved frequencies above $f_b/2$,

$$A_b(t) = B(t) - B_b(t) \quad \text{and} \quad \tilde{A}_b(f) = \tilde{B}(f) - \tilde{B}_b(f), \quad (11)$$

but there is no difference for frequencies below $f_b/2$.

Brick-wall filtering has two potential drawbacks. First, in the time domain, $B_b(t)$ will show oscillations or “ringing”

near discontinuities or abrupt changes in the original, unfiltered time series $B(t)$ (Bracewell, 1978, p. 209; Kanasewich, 1981, p. 240). This is the Gibbs phenomenon. It is usually illustrated in textbooks by convolving a sinc function with a Heaviside step function; since the sinc function has oscillatory tails, the output of the convolution is also, to some degree, oscillatory. While natural geomagnetic time series have no discontinuities, they do sometimes record abrupt variation, such as during storm sudden commencements. For this reason, the representation of a geomagnetic time series in terms of a brick-wall-truncated Fourier expansion will have ringing, which can be reasonably described as a Gibbs effect. But as we shall see, this does not significantly affect individual hourly samples of a brick-wall-filtered time series.

The second potential drawback of brick-wall filtering concerns its practical implementation. In comparison to the other average-sample types considered here, it is relatively more difficult to use brick-wall filtering to obtain hourly values. Accurate results require a long time series to be treated in the frequency domain. In contrast, simple boxcar averaging only requires one or two hours of data, and these can be treated in a straightforward way in the time domain; long time series are not needed to obtain accurate results. This simple point is especially relevant for institutes needing to promptly provide data for real-time operational applications.

3.3 Gaussian filtering

In the time domain, a Gaussian filter is defined as

$$\text{gaus}(t; t_\sigma) \equiv \exp\left(-\pi t^2/t_\sigma^2\right), \quad (12)$$

for which time-domain convolution is given by

$$B_\sigma(t) = \frac{1}{t_\sigma} B(t) * \text{gaus}(t; t_\sigma). \quad (13)$$

The Fourier transform of a Gaussian is another Gaussian,

$$\mathcal{F}\{\text{gaus}(t; t_\sigma)\} = \frac{1}{f_\sigma} \text{gaus}(f; f_\sigma) \quad (14)$$

(Bracewell, 1978, p. 386), where $f_\sigma = 1/t_\sigma$. In a qualitative sense, this time-frequency symmetry places the Gaussian filter midway between the two extremes of time-domain (frequency-domain) rectangle (sinc-function) filtering and sinc-function (brick-wall) filtering. The width of the Gaussian filter is specified by t_σ , and, in Sect. 5.7, we will locate an optimal value. We note that Intermagnet recommends a Gaussian filter for the production of 1-min averages, but they do not make any similar recommendation for hourly values.

3.4 Spot sampling

For a constant sampling frequency f_s with discrete sample times

$$t_j = t_1, t_2, t_3, \dots \quad (15)$$

that are equally-spaced with $t_{j+1} - t_j = t_s = 1/f_s$, we can define a Dirac comb function as a sequence of equally-spaced delta functions,

$$\text{comb}(t; t_s) \equiv t_s \sum_j \delta(t - jt_s) = \sum_j \delta(t/t_s - j) \quad (16)$$

(Bracewell, 1987, p. 77; Kanasewich, 1981, pp. 34, 110). With this, instantaneous or “spot” sampling of a time series can be described as a multiplication of the original, continuous time series by the Dirac comb,

$$B_s(t_j) = B(t) \text{comb}(t; t_s) = \dots B(-2t_s), B(-t_s), B(0), B(+t_s), B(+2t_s), \dots \quad (17)$$

(Bracewell, 1978, p. 78; Kanasewich, 1981, p. 35). Since the Fourier transform of a Dirac comb is another Dirac comb,

$$\mathcal{F}\{\text{comb}(t; t_s)\} = \frac{1}{f_s} \text{comb}(f; f_s), \quad (18)$$

in the frequency domain, convolution results in the superposition of an infinite sequence of replicas of $\tilde{B}(f)$, each one shifted by a multiple of the sampling frequency,

$$\begin{aligned} \mathcal{F}\{B_s(t_j)\} &= \frac{1}{f_s} \tilde{B}(f) * \text{comb}(f; f_s) \\ &= \sum_k \tilde{B}(f - kf_s). \end{aligned} \quad (19)$$

For a continuous time series defined across an infinite range of frequencies, it is impossible to fully resolve its frequency content with a discrete set of spot samples, even if we have an infinite number of them. This limitation is the result of “aliasing”, and it can be understood from consideration of Eq. (19). Harmonic amplitudes for frequencies below, for example, $f_s/2$ will have mapped onto them the amplitudes for frequencies above $f_s/2$, something that results in contamination of estimates of the frequency content of the original time series. More specifically, consider a single Fourier harmonic with frequency $0.3f_s$. With discrete sampling, it cannot be distinguished from a harmonic with frequency $0.7f_s$, or one with a frequency of $1.3f_s$, etc. There are an infinite number of aliases, each having a different frequency.

If a continuous time series contains only a compact range of frequencies, such that $\tilde{B}(f)$ is zero outside of the band $|f| < f_s/2$, then the replicas $\tilde{B}(f - kf_s)$ in Eq. (19) are non-overlapping. In this case, there will be no aliasing and Shannon’s theorem (Bracewell, 1978, Ch. 10; Kanasewich, 1981, p. 117) applies: sampling at twice Nyquist, $f_s = 2f_N$, is sufficient to make a complete reconstruction of the original, continuous time series. To see this, consider the frequency-limited version of Eq. (19),

$$\begin{aligned} \tilde{B}_s(f) &= \frac{1}{f_s} \left[\tilde{B}(f) * \text{comb}(f; f_s) \right] \text{rect}(f; f_s) \\ &= \sum_k \tilde{B}(f - kf_s) \text{rect}(f; f_s). \end{aligned} \quad (20)$$

If $\tilde{B}(f)$ has no frequency content above Nyquist, then only the $k = 0$ term contributes to the summation, in which case,

$$\tilde{B}_s(f) = \tilde{B}(f) \text{rect}(f; f_s) = \tilde{B}(f). \quad (21)$$

In the time domain, the dual of Eq. (20) is the Whittaker-Shannon “cardinal” formula (Bracewell, 1978, p. 413),

$$B_s(t) = \frac{1}{t_s} [B(t) \text{comb}(t; t_s)] * \text{sinc}(t; t_s). \quad (22)$$

The presence of the rectangle function in Eq. (20) makes clear the noteworthy property of Whittaker-Shannon interpolation: it does not introduce any spurious amplitudes with frequencies above Nyquist, and it does not change any amplitudes with frequencies below Nyquist. For these reasons, it is sometimes called an “ideal” interpolator. If $\tilde{B}(f)$ has no frequency content above Nyquist, then we can infer from Eq. (21) that

$$B_s(t) = \frac{1}{t_s} B(t) * \text{sinc}(t; t_s) = B(t). \quad (23)$$

A limited-frequency, continuous time series is, thus, reconstructed from spot samples by Whittaker-Shannon interpolation.

The situation of interest, here, is slightly messier. A continuous magnetic-field time series $B(t)$ has variation across a range of frequencies that is so broad that it might as well be considered infinite. But for a sampling given by (t_s, f_s) , we cannot resolve amplitudes with frequencies above Nyquist. That is, we are unable to resolve the following signal

$$\tilde{U}_s(f) = \tilde{B}(f) [1 - \text{rect}(f; f_s)], \quad (24)$$

$$U_s(t) = \frac{1}{t_s} B(t) * [\delta(t) - \text{sinc}(t; t_s)]. \quad (25)$$

Aliasing is given by

$$\tilde{S}_s(f) = \sum_{k \neq 0} \tilde{B}(f - kf_s) \text{rect}(f; f_s), \quad (26)$$

and, although non-standard, it is useful to view aliasing in the time domain,

$$\mathcal{F}^{-1}\{\tilde{S}_s(f)\} = S_s(t). \quad (27)$$

This has the spot values

$$S_s(t_j) = \dots S_s(-2t_s), S_s(-t_s), S_s(0), S_s(+t_s), S_s(+2t_s), \dots \quad (28)$$

which might be interpreted as “errors” in a frequency-limited representation of the original time series. For a qualitative discussion of aliasing in the context of geomagnetism, see Chapman and Bartels (1962, Ch. 16.14).

As with the amplitude differences defined by Eqs. (8) and (11), it is useful to define the residual differences between the original, continuous time series and the Whittaker-Shannon interpolation of spot samples,

$$R_s(t) = B(t) - B_s(t) \quad \text{and} \quad \tilde{R}_s(f) = \tilde{B}(f) - \tilde{B}_s(f), \quad (29)$$

where $B_s(f)$ and $B_s(t)$ are given by Eqs. (20) and (22). These residuals are equal to the difference between the unresolved, high-frequency signal and that of aliasing,

$$R_s(t) = U_s(t) - S_s(t) \quad \text{and} \quad \tilde{R}_s(f) = \tilde{U}_s(f) - \tilde{S}_s(f), \quad (30)$$

from Eqs. (24)–(27).

3.5 Spot sampling of a running average

Combining results, discrete average-sampling of a continuous time series is given by

$$B_{as}(t_j) = B_a(t) \text{comb}(t; t_s) \quad (31)$$

$$= \cdots B_a(-2t_s), B_a(-t_s), B_a(0), B_a(+t_s), B_a(+2t_s), \cdots$$

Its time-continuous Whittaker-Shannon interpolation is

$$B_{as}(t) = \frac{1}{t_s} [B_a(t) \text{comb}(t; t_s)] * \text{sinc}(t; t_s) \quad (32)$$

$$= \frac{1}{t_a t_s} [[B(t) * \text{rect}(t; t_a)] \text{comb}(t; t_s)] * \text{sinc}(t; t_s)$$

and

$$\tilde{B}_{as}(f) = \frac{1}{f_s} [\tilde{B}_a(f) * \text{comb}(f; f_s)] \text{rect}(f; f_s) \quad (33)$$

$$= \frac{1}{f_s} [[\tilde{B}(f) \text{sinc}(f; f_a)] * \text{comb}(f; f_s)] \text{rect}(f; f_s).$$

The frequency-domain residual difference is given by

$$\tilde{R}_{as}(f) = \tilde{B}(f) - \tilde{B}_{as}(f) \quad (34)$$

$$= \tilde{U}_s(f) + \tilde{L}_{as}(f) = \tilde{U}_s(f) + \tilde{A}_{as}(f) - \tilde{S}_{as}(f).$$

The unresolved, high-frequency signal is given by Eq. (24). The low-frequency residual,

$$\tilde{L}_{as}(f) = \tilde{A}_{as}(f) - \tilde{S}_{as}(f), \quad (35)$$

is the difference between amplitude distortion,

$$\tilde{A}_{as}(f) = \tilde{A}_a(f) \text{rect}(f; f_s), \quad (36)$$

where $\tilde{A}_a(f)$ is given by Eq. (8), and aliasing,

$$\tilde{S}_{as}(f) = \tilde{B}_{as}(f) - \tilde{B}_a(f) \text{rect}(f; f_s), \quad (37)$$

which should be compared with Eq. (26). Each of these amplitude distortion and aliasing terms can, of course, be expressed in the time domain as well.

The output of average-sampling depends on the chosen t_a and t_s (or, equivalently, f_a and f_s) and the details of the continuous time series itself. The standard practice of reporting 1-h-averages, once per hour, corresponds to $t_a = t_s$ ($f_a = f_s$). We have often heard it said that such average-samples are contaminated by aliasing, and that this might be remedied by, for example, increasing the averaging time from 1-h to 2-h, corresponding to $t_a = 2t_s$ ($2f_a = f_s$). But since there is considerable natural geomagnetic activity at periods below

2 h, heavier averaging will also result in significant amplitude distortion. Therefore, it is important to consider the combined effects of averaging and sampling on the continuous time series of interest: prominent Fourier amplitudes having periods that are much longer (shorter) than the averaging duration will not be (will be) significantly affected by amplitude distortion. If high-frequency Fourier amplitudes are relatively very small (very large), such as for what is often called a “red” (“blue”) spectrum, then aliasing might not be (could be) very important. See Kirchner (2005) for a discussion of the effects of averaging and sampling a time series having $1/f$ noise.

4 Numerical analysis

To decompose observatory data in terms of Fourier series, we first detrend each time series by subtracting a slowly changing, non-periodic trendline. The physical origin of this time-dependent trendline is the superposition of crustal magnetization beneath each observatory, which affects the average value of the trendline, and the time-varying main field sustained by the core’s dynamo, which affects the average value and the time dependence of the trendline. For each observatory, we approximate this internal-field time series by Chebyshev polynomials of the first kind (Press et al., 1992, “chebev”). There is precedence for this choice of basis function in geomagnetic analyses (e.g. Bloxham and Jackson, 1989), made because of the rapid rate with which Chebyshev expansion coefficients converge when approximating smooth functions (e.g. Conte and de Boor, 1980). A low-degree truncation of the Chebyshev expansion is fitted to the observatory data using a least-squares algorithm. Results are not, however, particularly sensitive to the truncation level (we choose one Chebyshev degree per 2 years of data), since the hourly samples are “high frequency” compared to the shortest characteristic timescale of the subtracted trendline.

We make transformations back and forth between the time and frequency domains by computer application of a fast-Fourier transform (Press et al., 1992, “realft”). This algorithm uses $N_B = 2^N$ data (amplitudes) in the time (frequency) domain, where N is an integer. The Fourier transformation of a discrete data set can be represented as

$$\mathcal{F}\{B(t_j)\} = \tilde{B}(f_k). \quad (38)$$

In the time domain, the discrete data

$$B(t_j) = B(t_1), B(t_2), B(t_3), \cdots \quad (39)$$

are assumed to be evenly-spaced, with a sampling interval $t_s = 1/f_s$. In the frequency domain, the amplitude pairs

$$[_s \tilde{B}(f_k), _c \tilde{B}(f_k)] = \quad (40)$$

$$[_s \tilde{B}(f_0), _c \tilde{B}(f_0)], [_s \tilde{B}(f_1), _c \tilde{B}(f_1)], [_s \tilde{B}(f_2), _c \tilde{B}(f_2)], \cdots$$

correspond to sine and cosine functions (denoted by the s and c prescripts), with the evenly-spaced discrete frequencies

$$f_k = kf_s/N_B = 0, f_s/N_B, 2f_s/N_B, \dots \quad (41)$$

As an approximation of continuous magnetic-field variation at each observatory, we use 1-min data. From them we synthesize approximate average-samples that are similar to historical (or possibly, prospective) reported observatory hourly values. But in order to facilitate comparison of results with the source 1-min data in both the time and frequency domains, instead of sampling the source 1-min time series once per hour, we sample them once every $64 = 2^6$ min. This means that for N_B 1-min data, the first $N_B/64$ Fourier amplitudes can be directly compared with the $N_B/64$ Fourier amplitudes from the “hourly” samples – they correspond to the same harmonics over the same total duration of time. For our purposes, this 64-min “hour” sampling period is sufficiently close to a normal 60-min hour.

Five different synthetic average-sample types, each constructed from the source minute data, are considered:

1. “Spot” samples, with no actual averaging,

$$B_s(t_j) = B(t_j). \quad (42)$$

2. 1-“hour” rectangle averages from $t_1 = 65$ min of data,

$$B_{1s}(t_j) = \frac{1}{t_1} \sum_{m=-32}^{m=+32} B(t_{j+m}) \text{rect}(t_{j+m}; t_1). \quad (43)$$

3. 2-“hour” rectangle averages from $t_2 = 129$ min of data,

$$B_{2s}(t_j) = \frac{1}{t_2} \sum_{m=-64}^{m=+64} B(t_{j+m}) \text{rect}(t_{j+m}; t_2). \quad (44)$$

4. Gaussian averages with $t_\sigma = 55$ min,

$$B_{\sigma s}(t_j) = \frac{1}{t_\sigma} \sum_{m=-\infty}^{m=+\infty} B(t_{j+m}) \text{gaus}(t_{j+m}; t_\sigma). \quad (45)$$

5. Brick-wall (sinc) averages,

$$B_{bs}(t_j) = \frac{1}{t_b} \sum_{m=-\infty}^{m=+\infty} B(t_{j+m}) \text{sinc}(t_{j+m}; t_b). \quad (46)$$

In each case, summation over 1-min data is denoted by the subscript m . The subscript j can, in principle, denote any sampling rate less than or equal to 1-min.

We will use both 1-min and 64-min sampling. With 1-min sampling, results correspond to “continuous” time series. For example, if the t_j are taken as minute sample times, then the corresponding spot samples are the same as the source 1-min data, $B_s(t_j) = B(t_j) \simeq B(t)$. With 64-min sampling, results correspond to discrete “hourly” samples of the time

series. Our choice, in Eqs. (43) and (44), to sum over 65 and 129 min of data, instead of 64 and 128, is motivated by the need to keep these average-samples symmetric in time and comparable to the other average-sample types. In practice, the summations in Eqs. (45) and (46) need to be taken over sufficiently large number of data to ensure convergence. However, it is easier to construct (46) by band-pass filtering in the frequency domain. In principle, t_b can be anything we want it to be, but it is usually set equal to twice the sampling interval, $t_b = 2t_s$, which is what we do.

5 Results

5.1 Frequency spectra of historical CLF data

The power spectral density for a time series $B(t)$ is defined by a “periodogram” (Kanasewich, 1981, Sect. 7.1),

$$P^B(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \left| \int_{-T/2}^{+T/2} B(t) e^{-2\pi i f t} dt \right|^2, \quad (47)$$

and the display of $P^B(f)$ in a graph as a function of frequency f is a “spectrum”. The discrete power spectral density, for each unit of frequency, is given by

$$P^B(f_k) = \frac{2}{N_B^2} \left\{ |s \tilde{B}(f_k)|^2 + |c \tilde{B}(f_k)|^2 \right\}. \quad (48)$$

In Fig. 1a we show power spectra $P^H(f)$ for two 18-year durations of CLF horizontal-intensity hourly values: 1954.0–1972.0 (hourly spot measurements) and 1972.0–1990.0 (1-h averages). Regular solar-quiet variation is seen as prominent diurnal peaks with frequencies of 1/d, 2/d, etc. (e.g. Olsen, 2007). Magnetic storms and disturbance are seen as a broad wash of energy across the entire range of frequencies, with a gradual “red-spectrum” decrease in energy with increasing frequency.

Important for this study are the differences in the spectra between the older pre-1972.0 data, when reported hourly values were spot measurements, and the newer post-1972.0 data, when reported hourly values were 1-h average-samples. The spectral power for spot measurements is higher at high frequencies than that for 1-h average-samples. Since spectral differences might simply be reflective of different levels of activity, these observations, on their own, are not sufficient to prove that different average-sample types (spot and average) suffer from different amounts of amplitude distortion and aliasing. But what we see in Fig. 1a is enough to motivate further investigation. We will return, at the end of Sect. 5.3, to discuss Fig. 1b.

5.2 Example of average-sampling during a storm

Using 1-min CLF- H source data, 1991.0–2009.0, and formulas (42) and (43), we generated synthetic “spot” samples, “continuous” running averages, and discrete “hourly”

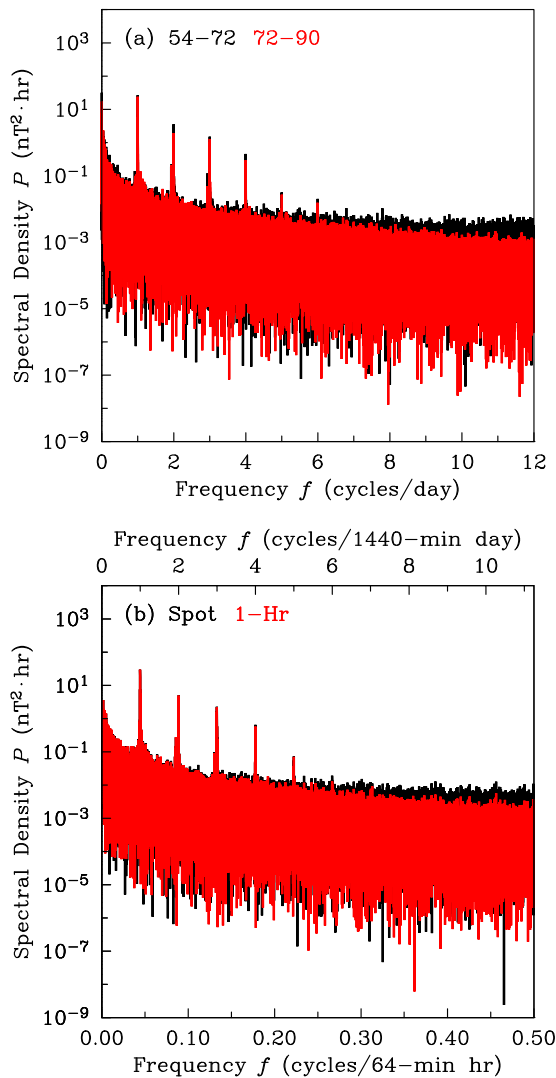


Fig. 1. Comparison of (a) power spectral density $P^H(f)$ as a function of frequency f for historical hourly CLF- H data for years when instantaneous “spot” values were reported, 1954.0–1972.0 (black), and when 1-h average-samples were reported, 1972.0–1990.0 (red); (b) synthetic spot samples (black) and 1-h average-samples (red), each constructed from 1-min CLF- H data, 1991.0–2009.0.

average-samples. In Fig. 2 we show results for a period of time recording a large magnetic storm (October 2003). As expected, we see in Fig. 2a that spot samples $H_s(t_j)$ sometimes record the extreme amplitudes of transient variation, such as for the dramatic initial phase of this storm (Day 302.3) and during the storm’s second main phase (Day 303.8–304.1), while at other times transient variation is missed, such as during the hours immediately following the initial phase (Day 302.4). Magnetic-field variation at frequencies higher than Nyquist is not well represented by spot sampling, and because of this there is substantial aliasing, $S_s(t_j)$ Eq. (28). The signals of continuous interpolations of

the spot samples, $H_s(t)$ Eq. (22), and the discrete aliases, $S_s(t)$ Eq. (27), show Gibbs ringing. This tends to be correlated with spot samples that happen to fall on the extremes of the 1-min variation. Otherwise, before and after periods of rapid variation, ringing tends to occur in between spot values. The latter is a result of the fact that for $t_s = 1$ h the sinc function has zero crossings at integer hours, a point we will examine again in Sect. 5.8.

In Fig. 2b we see that continuous, running averaging gives a smooth representation of the original time series, but the extreme amplitudes of storm-time variation are smoothed out. We have described this as amplitude distortion, $A_1(t)$ Eq. (8). Discrete sampling of the averaged time series has a continuous signal, $H_{1s}(t)$ Eq. (32), with less Gibbs ringing than that for spot sampling, and a low-frequency residual, $L_{1s}(t)$ Eq. (35), that is, on average, less than aliasing of spot sampling.

5.3 Frequency spectra of different average-samples

Using, again, 1-min CLF- H source data, 1991.0–2009.0, we used formulas (42) to (46) to generate a variety of average-sample types. In Fig. 3a we see that the hourly spot samples have a spectrum $P_s^H(f)$ that is generally like “continuous variation” that is approximated by 1-min data $P^H(f)$. However, at high frequencies, near the Nyquist frequency (~ 0.5 cycles/h), the energy of the hourly spot spectrum is higher than that for the 1-min data. This lifting of the high-frequency end of the hourly spectrum is entirely due to aliasing, $\tilde{S}_s(f)$ Eq. (26). In Fig. 3b we show the spectrum of aliasing $P_s^S(f)$, which can be recognized as the reflection of the 1-min spectrum through the Nyquist frequency. Because the 1-min spectrum is “red”, reflection through Nyquist results in an aliasing spectrum that is “blue”, and for this reason, aliasing is most easily seen in Fig. 3a for frequencies immediately below Nyquist.

More interesting are the spectra for 1-h average-samples, Fig. 3c. The spectrum $P_1^H(f)$ of continuous running averages is depressed at high frequencies compared to the 1-min spectrum $P^H(f)$ by an amount equal to the square of the amplitude-distorting, frequency-response factor, $\text{sinc}^2(f; f_1)$ Eq. (7). While the spectrum of discretely-sampled 1-h average values $P_{1s}^H(f)$ is slightly lower than that of the 1-min data, it is also slightly higher than the spectrum for continuous running averages. Again, this difference is due to aliasing, $\tilde{S}_{1s}(f)$ Eq. (37). In Fig. 3d we show the low-frequency residual spectrum, $P_{1s}^L(f)$ Eq. (35), and the aliasing spectrum, $P_{1s}^S(f)$ Eq. (37). Energy in the residual spectra increases with frequency (“blue”), and the integrated energy up to the Nyquist frequency is evidently less than that seen for instantaneous, spot samples Fig. 3b, an observation consistent with observations made of the time series in Fig. 2. Furthermore, while aliasing accounts for a large portion of the low-frequency residuals, it does not account for the relatively prominent residual differences seen in periodic diurnal

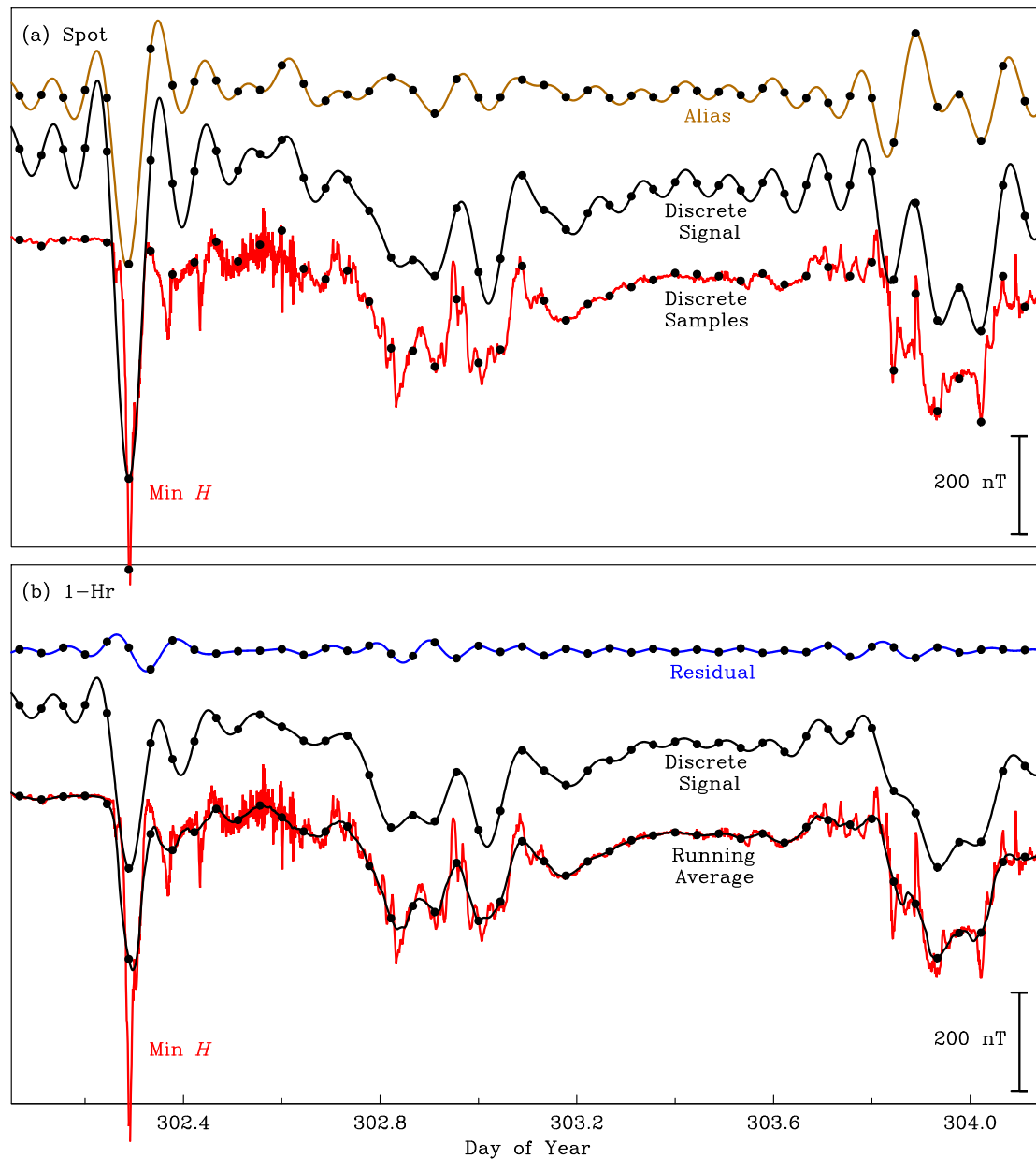


Fig. 2. Data from CLF recording 2 days of horizontal-intensity magnetic-field variation during the Halloween storm of October 2003: 1-min data $H(t)$ (red), running averages $H_a(t)$ (black), signal of discrete samples $H_{as}(t)$ (black), alias $S_s(t)$ (brown), low-frequency residual $L_{as}(t)$ (blue), and discrete samples (black dots), each for **(a)** instantaneous, spot samples $H_s(t_j)$ and **(b)** 1-h average-samples $H_{1s}(t_j)$.

terms – those differences are due to amplitude distortion. Similar conclusions can be drawn from the spectra for Gaussian average-samples Fig. 3g, h.

For 2-h running averages, we see in Fig. 3e that the energy in the spectrum $P_2^H(f)$ is substantially depressed for frequencies near Nyquist, and, indeed, the spectrum for hourly 2-h average-samples $P_{2s}^H(f)$ is also substantially depressed at high frequencies. In Fig. 3f we see that aliasing spectrum $P_2^S(f)$ is lower than that for 1-h Fig. 3d or Gaussian Fig. 3h average-samples, but the residual spectrum $P_2^L(f)$ is higher.

Brick-wall results, $P_\sigma^H(f)$ Fig. 3i, j, are a special case; this type of average-sample has perfect frequency response below Nyquist, and therefore no low-frequency residuals at all. The lesson we can take away from these comparisons is that there is a trade-off between amplitude distortion and aliasing.

Returning, now, to the power spectra of the historical CLF data, in Fig. 1b we show the spectra for hourly spot samples $P_s^H(f)$ and 1-h average-samples $P_{1s}^H(f)$. In each case, these have been constructed from 1-min data covering an 18-year (1991.0–2009.0) period of time that is equal in duration

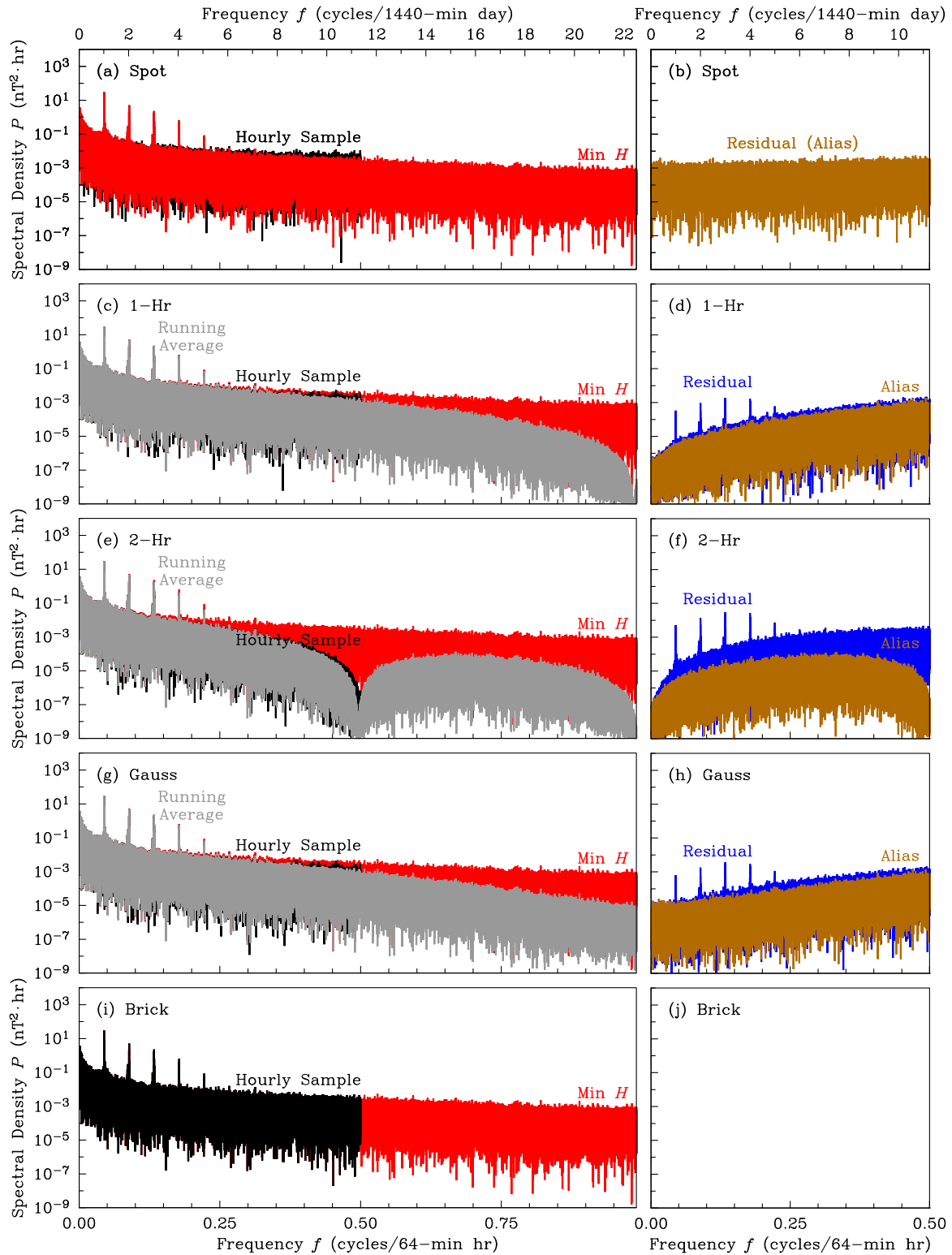


Fig. 3. Power spectral density $P^H(f)$ as a function of frequency f , for CLF- H 1991.0–2009.0: (a, b) spot samples, (c, d) 1-h, (e, f) 2-h, (g, h) Gaussian, and (i, j) brick-wall average-samples. Shown are spectral densities of the source 1-min data $P^H(f)$ (red), running averages $P_a^H(f)$ (gray), discrete hourly average-samples $P_{as}^H(f)$ (black), low-frequency residuals $P_{as}^L(f)$ (blue), and aliasing $P_{as}^S(f)$ (brown).

Table 2. Statistics of discrete 64-min “hourly” average-samples for 1998.0–2009.0.

Lat.	Obs.		Ave. Abs. Dev. δ			Standard Dev. σ		
			H (nT)	D ($^\circ$)	Z (nT)	H (nT)	D ($^\circ$)	Z (nT)
High	BRW	Spot	75.8	14.77	50.6	122.8	26.16	83.7
		1-h	69.8	11.82	48.6	107.1	18.87	79.5
		2-h	66.3	11.00	46.8	98.1	16.89	75.8
		Gauss	69.1	11.66	48.3	105.7	18.52	79.0
		Brick	71.9	12.23	49.5	110.6	19.42	81.1
Medium	CLF	Spot	14.5	2.32	7.8	21.6	3.23	11.8
		1-h	14.1	2.24	7.8	20.9	3.10	11.7
		2-h	13.8	2.17	7.7	20.4	2.97	11.4
		Gauss	14.1	2.23	7.8	20.8	3.08	11.6
		Brick	14.3	2.28	7.8	21.1	3.14	11.7
Low	HUA	Spot	43.4	8.94	6.3	58.6	16.90	8.6
		1-h	42.9	8.87	6.2	57.5	16.81	8.5
		2-h	42.1	8.78	6.1	56.1	16.68	8.2
		Gauss	42.8	8.86	6.2	57.3	16.79	8.4
		Brick	43.2	8.91	6.3	57.9	16.86	8.6

to the two periods of time shown in Fig. 1a for the historical data (spot samples 1954.0–1972.0, 1-h average-samples 1972.0–1990.0). From the considerable similarity between the spectra seen in Figs. 1a, b and 3a, b, we conclude that the spectral differences in the historical CLF, seen in Fig. 1a, are almost entirely due to different averaging and sampling methods.

5.4 Statistical moments of the average-samples

As a statistical summary of magnetic-field variability recorded by the various hourly average-sample types, we calculate average-absolute deviations

$$\delta^B = \frac{1}{N_B} \sum_j |B(t_j)|, \quad (49)$$

and standard deviations

$$\sigma^B = \left[\frac{1}{N_B} \sum_j |B(t_j)|^2 \right]^{\frac{1}{2}}, \quad (50)$$

each defined here for zero-mean data; recall from Sect. 4 that we have subtracted a slowly-varying trendline. By Parseval’s theorem, the variance of the time series equals the total power integrated over all frequencies,

$$\frac{1}{N_B} \sum_j |B(t_j)|^2 = \sum_k P^B(f_k) \quad (51)$$

(Lee, 1960, p. 11; Bracewell, 1978, p. 112). Therefore, results for statistical variance, Eq. (50), can be interpreted in terms of the integral of the spectral power in the frequency-domain, Eq. (48).

In Table 2 we list statistical moments for horizontal intensity H , declination D and vertical intensity Z , from high (BRW), medium (CLF), and low-latitude (HUA) observatories. Hourly spot measurements are an unbiased representation, in a statistical sense, of the amplitude range of magnetic-field variation. It is, therefore, useful to compare the moments of spot-sample measurements with those of the other average-samples types. Without exception, brick-wall (2-h) average-samples have δ and σ values closest to (furthest from) those of spot samples. Still, it is noteworthy that results for 1-h average-samples are relatively similar to those of both the Gaussian and brick-wall average-samples.

5.5 Low-frequency residual moments

In Table 3 we list the statistical moments of the low-frequency residuals, $L_{as}(t)$ Eq. (35). Generally speaking, the smaller the size of this residual, the better the representation of continuous magnetic-field variation for frequencies below Nyquist. Low-frequency residuals for spot samples are the largest, something that is, once again, entirely due to aliasing. Those for 2-h average-samples are the next largest, something that is primarily due to amplitude distortion. For 1-h and Gaussian average-samples, low-frequency residual differences are of similar size. There are no residuals for the brick-wall average-samples. For all average-sample types, even spot samples, the Z average absolute and standard deviations for medium latitudes (CLF) and low latitudes (HUA) are less than or equal to the 1.0 nT resolution used in historical observatory yearbooks.

Table 3. Statistics of low-frequency residuals $L_{\text{as}}(t)$ for 1998.0–2009.0.

Lat.	Obs.		Ave. Abs. Dev. δ			Standard Dev. σ		
			H (nT)	D (μ)	Z (nT)	H (nT)	D (μ)	Z (nT)
High	BRW	Spot	29.2	8.71	11.2	53.7	17.38	20.7
		1-h	11.4	2.65	5.0	17.7	4.24	7.8
		2-h	18.8	3.82	9.2	31.6	6.36	14.9
		Gauss	11.3	2.51	5.2	17.7	3.94	8.0
		Brick	0.0	0.00	0.0	0.0	0.00	0.0
Medium	CLF	Spot	2.1	0.39	0.4	4.2	0.75	1.0
		1-h	1.0	0.17	0.2	1.6	0.28	0.4
		2-h	1.8	0.31	0.6	3.0	0.50	0.9
		Gauss	1.0	0.17	0.3	1.6	0.28	0.4
		Brick	0.0	0.00	0.0	0.0	0.00	0.0
Low	HUA	Spot	4.3	0.55	0.6	8.7	1.31	0.9
		1-h	1.9	0.25	0.3	3.1	0.50	0.4
		2-h	3.4	0.50	0.7	5.5	0.97	0.9
		Gauss	1.9	0.26	0.3	3.0	0.52	0.4
		Brick	0.0	0.00	0.0	0.0	0.00	0.0

Table 4. Statistics of discrete 685×64 -min “monthly” differences $\epsilon_{\star m}(t_j)$ for 1998.0–2009.0.

Lat.	Obs.		Ave. Abs. Dev. δ			Standard Dev. σ		
			H (nT)	D (μ)	Z (nT)	H (nT)	D (μ)	Z (nT)
High	BRW	Spot	1.6	0.55	0.6	2.0	0.73	0.8
		1-h	0.0	0.00	0.0	0.0	0.00	0.0
		2-h	0.0	0.01	0.0	0.1	0.02	0.1
		Gauss	0.1	0.04	0.1	0.1	0.05	0.1
		Brick	0.0	0.01	0.0	0.0	0.01	0.0
Medium	CLF	Spot	0.1	0.02	0.0	0.1	0.02	0.0
		1-h	0.0	0.00	0.0	0.0	0.00	0.0
		2-h	0.0	0.00	0.0	0.0	0.00	0.0
		Gauss	0.0	0.00	0.0	0.0	0.00	0.0
		Brick	0.0	0.00	0.0	0.0	0.00	0.0
Low	HUA	Spot	0.2	0.03	0.0	0.3	0.05	0.0
		1-h	0.0	0.00	0.0	0.0	0.00	0.0
		2-h	0.0	0.01	0.0	0.1	0.00	0.0
		Gauss	0.0	0.01	0.0	0.0	0.02	0.0
		Brick	0.0	0.01	0.0	0.0	0.01	0.0

5.6 Monthly and annual means

Monthly and annual observatory means are used for secular variation studies, and so it is useful to evaluate the effects of using different types of hourly average-samples in their construction. For simplicity, we define “monthly” means as the average of geomagnetic-vector components over $t_m = 685$ “hours”, each being 64 min in length,

$$B_{\star m}(t_j) = \frac{1}{t_m} \sum_{h=-342}^{h=+342} B_{\star}(t_{j+h}), \quad (52)$$

with \star denoting any chosen hourly average-sample type from Eq. (42) to Eq. (46). The difference between a monthly mean of hourly average-samples and a monthly mean of 1-min data is

$$\epsilon_{\star m}(t_j) = B_{\star m}(t_j) - B_{1\text{sm}}(t_j). \quad (53)$$

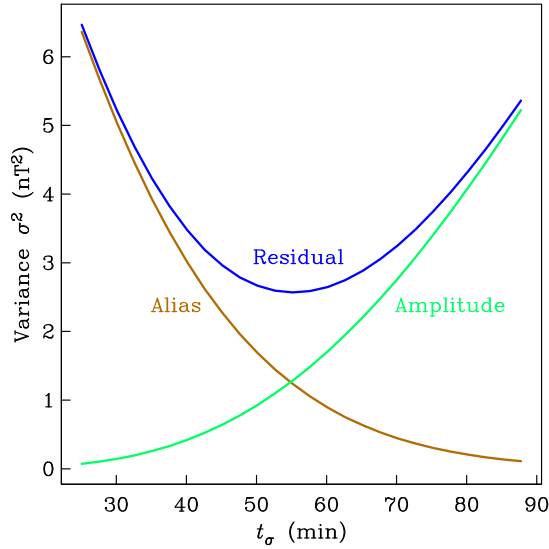


Fig. 4. Trade-off diagram, for CLF-*H* 1998.0–2009.0, showing variance of amplitude distortion $A_{\sigma s}(t)$ (green), aliasing $S_{\sigma s}(t)$ (brown), and low-frequency residuals $L_{\sigma s}(f)$ (blue) for Gaussian average-samples over a range of averaging times t_{σ} .

In Table 4 we list the statistical moments of $\epsilon_{*m}(t)$. For low and medium latitudes, average differences are less than the 1.0 nT resolution in the monthly-mean database of Chulliat and Telali (2007). As might be expected, differences are largest for spot samples from high-latitude observatories, but even these differences are small compared to the 5.0 nT standard for absolute accuracy established by Intermagnet for modern 1-min data. Corresponding differences for annual means will be even smaller, by about a factor of $1/\sqrt{12}$. The slightly larger average differences for Gaussian average-samples are mostly due to end-point differences, with the tails of each hourly Gaussian filter extending outside of each monthly boxcar averaging window. In analyzing secular variation, it is probably acceptable to use any type of hourly average-sample to construct monthly and annual means.

5.7 Optimum Gaussian filter

Our choice of a Gaussian filter with $t_{\sigma} = 55$ min minimizes the variance of the low-frequency residuals, $L_{\sigma s}(f)$ Eqs. (35) and (50). In Fig. 4 we show this quantity, along with the variance of amplitude distortion $A_{\sigma s}(t)$ and aliasing $S_{\sigma s}(t)$, as function of t_{σ} . Amplitude distortion (aliasing) increases (decreases) with increasing t_{σ} , but there is an inflection point at $t_{\sigma} = 55$ min where the residuals are minimized. This method of selecting an optimal averaging width is a simple example of what is sometimes called “filter design”.

It is tempting to try to relate the optimized Gaussian filter to a boxcar average of a certain averaging width. While we hesitate to read too much into such comparisons, we have already observed from Table 3 that some of the properties

of 1-h average-samples are similar to those of the optimum Gaussian filter. The small statistical differences between these two average-sample types are probably not very significant for most uses of observatory hourly values. Of course, the redeeming quality of 1-h average-samples is that they are easy to calculate, even easier than Gaussian average-samples.

5.8 Autocorrelation and avoiding the Gibbs effect

While it is a standard practice to use power spectra to check for periodic signals and to display Fourier amplitudes in the frequency domain, in the time domain the corresponding quantity is autocorrelation. For a real function, such as an observatory magnetic-field time series, autocorrelation is defined by the integral

$$C^B(\tau) = \lim_{\Phi \rightarrow \infty} \frac{1}{\Phi} \int_{-\Phi/2}^{+\Phi/2} B(\phi)B(\phi + \tau)d\phi, \quad (54)$$

(Press et al., 1992, “correl”), where τ represents a relative time lag. By the Wiener-Khinchine or “auto-correlation” theorem (Lee, 1960, p. 11; Bracewell, 1978, p. 115), an autocorrelation is related to a power spectrum by Fourier transformation,

$$\mathcal{F}\{C^B(\tau)\} = P^B(f). \quad (55)$$

Despite this duality, some properties of a time series are more easily seen in power spectra, while others are more easily seen in autocorrelation.

In Fig. 5 we show autocorrelations for the source 1-min data $C^H(\tau)$, continuous running averages $C_a^H(\tau)$, and average-samples $C_{as}^H(\tau)$ for CLF-*H* 1991.0–2009.0. Prominent solar-quiet diurnal peaks are seen with time lags of 1d, 2d, etc.; these correspond to the diurnal peaks in frequency of $1/d$ seen in the power spectra of Fig. 3. For very short time lags, less than the averaging duration, autocorrelations $C_a^H(\tau)$ are flat. These results are all to be expected. What is perhaps more interesting are the autocorrelations for low-frequency residuals, $C_a^L(\tau)$ Eq. (35). For spot sampling, residuals are entirely due to aliasing, and since aliasing residuals tend to be larger than those for other average-sample types, the autocorrelation of spot-sampling residuals has a larger zero-time-lag base value. For all average-sample types, other than the brick-wall, there is significant Gibbs-effect ringing out to time lags of several hours. Close inspection of Fig. 5 shows that the zero crossings of $C_a^L(\tau)$ occur, very nearly, at integer hours (64-min hours). This is consistent with the observation we made in Sect. 5.2 that hourly average-samples avoid most Gibbs ringing by straddling the zero crossings of the sinc function.

5.9 Correlation and proportionality of average-samples

As we have remarked, brick-wall average-samples have no low-frequency residuals, while the other average-sample

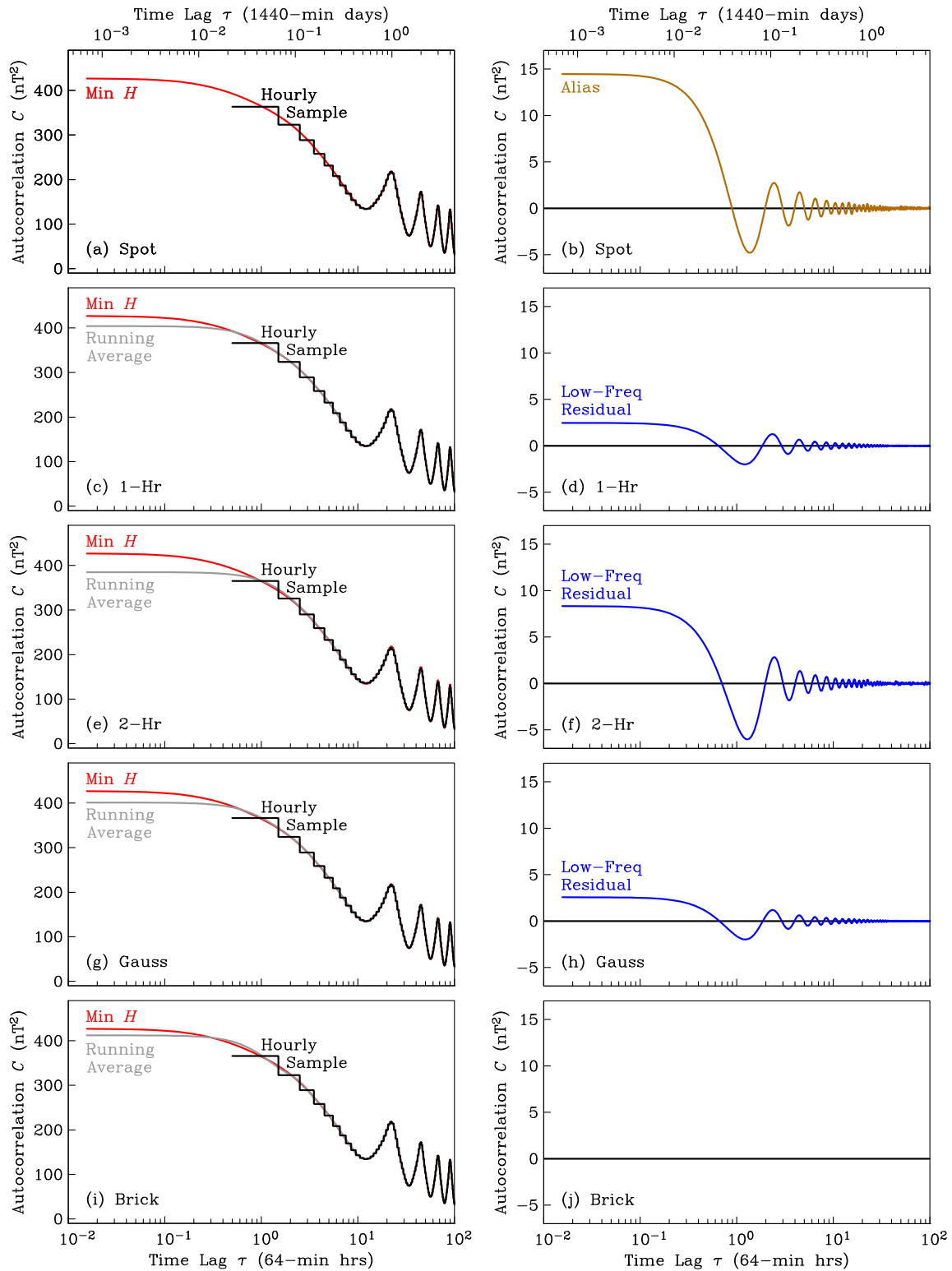


Fig. 5. Autocorrelation $C^H(\tau)$ as a function of time lag τ , for CLF- H 1991.0–2009.0: (a, b) spot (c, d) 1-h, (e, f) 2-h, (g, h) Gaussian, and (i, j) brick-wall average-samples. Shown are autocorrelation of the source 1-min data $C^H(\tau)$ (red), running averages $C_a^H(\tau)$ (gray), discrete hourly average-samples $C_{as}^H(\tau_j)$ (black), low-frequency residuals $C_a^L(\tau)$ (blue), and aliasing $C_s^S(\tau)$ (brown).

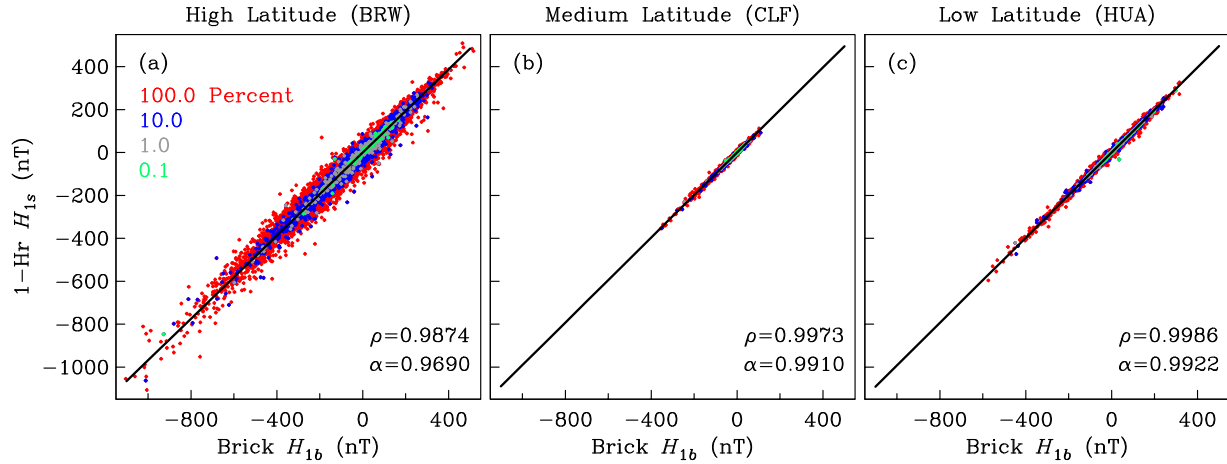


Fig. 6. Scatter plots of discrete 1-h $H_{1s}(t_j)$ versus brick-wall average-samples $H_{bs}(t_j)$ for (a) high latitude (BRW), (b) medium latitude (CLF), and (c) low latitude (HUA). Also given, for each case, are correlation coefficients ρ and proportionalities α (shown as a fitted line). Data used are for 1998.0–2009.0, and densities are for 100.0, 10.0, 1.0, and 0.1% of the data.

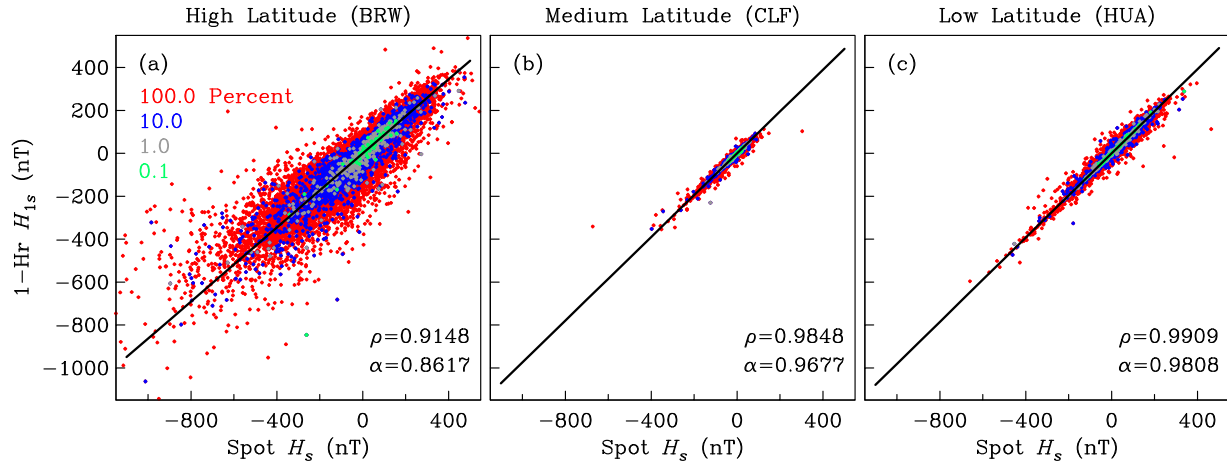


Fig. 7. Scatter plots of discrete 1-h average-samples $H_{1s}(t_j)$ versus spot samples $H_s(t_j)$ for (a) high latitude (BRW), (b) medium latitude (CLF), and (c) low latitude (HUA). Also given, for each case, are correlation coefficients ρ and proportionalities α (shown as a fitted line). Data used are for 1998.0–2009.0, and densities are for 100.0, 10.0, 1.0, and 0.1% of the data.

types have some combination of amplitude distortion and aliasing. It is useful, therefore, to compare linear correlations ρ between brick-wall and other average-sample types,

$$\rho = \frac{1}{N_B} \sum_j \frac{B_{as}(t_j) B_{bs}(t_j)}{\sigma_{as}^B \sigma_{bs}^B}, \quad (56)$$

(Press et al., 1992, “pearsn”), defined here for zero-mean data; recall from Sect. 4 that we have subtracted a slowly-varying trendline. A correlation close to unity means that the relative amplitude, phase, and frequency of continuous magnetic-field variation below Nyquist are being accurately recorded. In Fig. 6 we show scatter plots for brick-wall and 1-h average-samples. In Table 5 we list correlations for each magnetic-field component and for each observatory. In general, spot and 2-h average-samples have the lowest correla-

tions with brick-wall average-samples, while 1-h and Gaussian average-samples are very highly correlated with brick-wall average-samples.

Since spot measurements are a statistically unbiased sampling of the amplitude range of magnetic-field variation, it is useful to compare their proportionalities α with the other average-samples,

$$B_{as}(t_j) \simeq \alpha B_s(t_j), \quad (57)$$

where each α are estimated with a least-squares algorithm (Press et al., 1992, “fitxy”). A proportionality close to unity means that the average-sample type is accurately recording the absolute amplitude range of magnetic-field variation. In Fig. 7 we show scatter plots for spot and 1-h average-samples. In Table 5 we list proportionalities for each

Table 5. Correlations with brick-wall average samples and proportionalities with spot samples, 1998.0–2009.0.

Lat.	Obs.		Corr. with brick ρ			Prop. with spot α		
			H	D	Z	H	D	Z
High	BRW	Spot	0.9001	0.7475	0.9690	1.0000	1.0000	1.0000
		1-h	0.9874	0.9759	0.9955	0.8617	0.6565	0.9485
		2-h	0.9615	0.9480	0.9841	0.7687	0.5393	0.9013
		Gauss	0.9878	0.9795	0.9954	0.8597	0.6496	0.9419
		Brick	1.0000	1.0000	1.0000	0.8902	0.6745	0.9685
Medium	CLF	Spot	0.9818	0.9728	0.9963	1.0000	1.0000	1.0000
		1-h	0.9973	0.9962	0.9994	0.9677	0.9573	0.9862
		2-h	0.9906	0.9881	0.9972	0.9427	0.9142	0.9669
		Gauss	0.9973	0.9962	0.9994	0.9649	0.9514	0.9837
		Brick	1.0000	1.0000	1.0000	0.9773	0.9713	0.9936
Low	HUA	Spot	0.9891	0.9989	0.9944	1.0000	1.0000	1.0000
		1-h	0.9986	0.9998	0.9989	0.9808	0.9946	0.9837
		2-h	0.9959	0.9994	0.9948	0.9578	0.9873	0.9563
		Gauss	0.9987	0.9998	0.9988	0.9775	0.9935	0.9799
		Brick	1.0000	1.0000	1.0000	0.9888	0.9975	0.9947

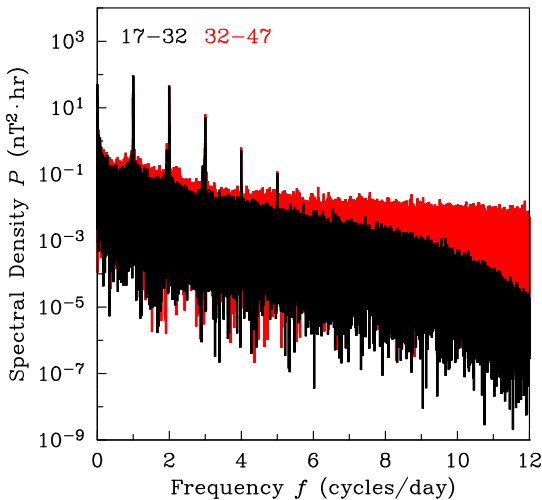


Fig. 8. Comparison of power spectral density $P^H(f)$ as a function of frequency f for historical hourly ESK- H data obtained from WDC-CE, 1917.0–1932.0 (black) and 1932.0–1947.0 (red).

magnetic-field component and for each observatory. Without exception, brick-wall (2-h) average-samples are the most (least) directly proportional to spot samples. Results for 1-h average-samples are relatively similar to those of both Gaussian and brick-wall average-samples.

5.10 ESK hourly values at the WDCs

In Sect. 2 we mentioned that the pre-1932.0 hourly WDC-CE ESK data are 2-h average-samples, with each datum formed by averaging two adjacent 1-h average-samples from

the original observatory yearbooks. In Fig. 8 we show frequency-domain power spectra of historical ESK- H data obtained from WDC-CE, separately for the years 1917.0–1932.0 and 1932.0–1947.0. As a result of 2-h averaging, the frequency content of the data for 1917.0–1932.0 has been dramatically altered. Since many researchers assume that data obtained from a WDC are those that the observatories originally reported, it would be desirable to return the WDC-CE holdings of the ESK data to their original form, possibly by using the digital data held at Kyoto. Otherwise, it would be useful for the data to be flagged as having been changed.

Note added after journal review: we are happy to have learned that the pre-1932.0 hourly WDC-CE ESK data have been restored to their original yearbook values (S. Macmillan, personal communication, 2010).

6 Conclusions

In analyses complementary to this one (e.g. Love, 2009; Marsal and Curto, 2009), the accuracy of hourly average-samples constructed from digital 1-min data having missing values has been investigated, and corresponding errors were estimated. Here we have considered a more fundamental issue, the effects of averaging and sampling in the construction of observatory hourly values. Of the average-sample types most commonly used, spot samples are a statistically-unbiased representation of the amplitude range of geomagnetic-field variation, but as a representation of continuous variation over time, spot values are heavily contaminated by aliasing. On the other hand, 1-h average-samples are a statistically-biased representation of geomagnetic-field values, but as a representation of continuous variation below

Nyquist, 1-h average-samples from medium and low-latitude observatories are only slightly contaminated by amplitude distortion and aliasing – their average low-frequency residuals are less than the 5.0 nT absolute accuracy level established by Intermagnet as a modern operational standard for observatories. High-latitude 1-h average-samples have more amplitude distortion and aliasing, but the natural magnetic “signal” is also larger at high latitudes. 1-h average-samples constructed using an optimal Gaussian filter are not significantly more accurate than simple 1-h average values. Therefore, we recommend the continuation of the standard practice by observatories and World Data Centers for reporting simple 1-h-average hourly values. Of course, this does not preclude the production of other data-based products, including other types of hourly values, but maintaining continuity with the historical 1-h-average time series should be recognized as an important priority.

Acknowledgements. The results presented here rely on data collected at the Barrow, Chambon la Forêt, Eskdalemuir, and Huan-cayo observatories. The Barrow observatory is operated by the USGS Geomagnetism Program. We thank the Institut de Physique du Globe de Paris (France), the British Geological Survey (Great Britain), and the Instituto Geofísico (Perú) for continuing to support the operation of their observatories. We acknowledge Intermagnet (www.intermagnet.org) for facilitating the dissemination of magnetic observatory data and for promoting high standards of magnetic observatory practice. We thank the Edinburgh and Kyoto World Data Centers for archiving observatory data and making them readily available to the scientific community. We thank the Boulder World Data Center (NGDC) for use of their observatory yearbooks. We thank P. S. Earle, J. W. Kirchner, K. Mursula, and L. Svalgaard for useful conversation/communication. We thank A. Chulliat, C. A. Finn, V. Lesur, and S. Macmillan, for reviewing a draft manuscript.

Topical Editor I. A. Daglis thanks S. Macmillan, A. Chulliat, and another anonymous referee for their help in evaluating this paper.

References

- Barracough, D. R., Clark, T. D. G., Cowley, S. W. H., Hibberd, F. H., Hide, R., Kerridge, D. J., Lowes, F. J., Malin, S. R. C., Murphy, T., Rishbeth, H., Runcorn, S. K., Soffel, H. C., Stewart, D. N., Stuart, W. F., Whaler, K. A., and Winch, D. E.: 150 years of magnetic observatories: Recent researches on world data, *Surv. Geophys.*, 13, 47–88, 1992.
- Bendat, J. S. and Piersol, A. G.: *Random Data: Analysis and Measurement Procedures*, John Wiley & Sons, New York, NY, 2000.
- Bloxham, J. and Jackson, A.: Simultaneous stochastic inversion for geomagnetic main field and secular variation 2. 1820–1980, *J. Geophys. Res.*, 94, 15753–15769, 1989.
- Bracewell, R. N.: *The Fourier Transform and its Applications*, McGraw-Hill Book Company, New York, NY, 1978.
- Brooke, C.: On the automatic registration of magnetometers, and other meteorological instruments, by photography, *Phil. Trans. R. Soc. Lond.*, 137, 69–77, 1847.
- Campbell, W. H.: The regular geomagnetic-field variations during quiet solar conditions, in: *Geomagnetism*, Volume 3, edited by: Jacobs, J. A., pp. 385–460, Academic Press, London, UK, 1989.
- Chapman, S. and Bartels, J.: *Geomagnetism*, Volumes 1 & 2, Oxford Univ. Press, London, UK, 1962.
- Chulliat, A. and Telali, K.: World monthly means database project, in: *Proceedings of the XIIth IAGA Workshop on Geomagnetic Instruments, Data Acquisition, and Processing*, edited by: Reda, J., pp. 268–274, *Publs. Inst. Geophys. Pol. Acad. Sc.*, 2007.
- Conte, S. D. and de Boor, C.: *Elementary Numerical Analysis*, McGraw-Hill Book Company, New York, NY, 1980.
- Courtillot, V. and Le Mouél, J. L.: Time variations of the Earth’s magnetic field: From daily to secular, *Ann. Rev. Earth Planet. Sci.*, 16, 389–476, 1988.
- Egbert, G. D., Booker, J. R., and Schultz, A.: Very long period magnetotellurics at Tucson Observatory: Estimation of impedances, *J. Geophys. Res.*, 97, 15113–15128, 1992.
- Fouassier, D. and Chulliat, A.: Extending backwards to 1883 the French magnetic hourly data series, in: *Proceedings of the XIIth IAGA Workshop on Geomagnetic Instruments, Data Acquisition, and Processing*, edited by: Love, J. J., pp. 86–94, US Geological Survey, 2009.
- Hazard, D. L.: *Results of Observations Made at the United States Coast and Geodetic Survey Magnetic Observatory at Cheltenham, Maryland, 1915 and 1916*, Govt. Printing Office, Washington, D.C., 1918.
- Jankowski, J. and Sucksdorff, C.: *Guide for Magnetic Measurements and Observatory Practice*, IAGA, Warsaw, Poland, 1996.
- Kanasewich, E. R.: *Time Sequence Analysis in Geophysics*, Univ. Alberta Press, Alberta, Canada, 1981.
- Karinen, A. and Mursula, K.: A new reconstruction of the D_{st} index for 1932–2002, *Ann. Geophys.*, 23, 475–485, doi:10.5194/angeo-23-475-2005, 2005.
- Kerridge, D. J.: INTERMAGNET: Worldwide near-real-time geomagnetic observatory data, *Proc. ESA Space Weather Workshop, ESTEC, Noordwijk, The Netherlands*, <http://www.esa-spaceweather.net/spweather/workshops/workshops.html>, 2001.
- Kirchner, J. W.: Aliasing in $1/f^\alpha$ noise spectra: Origins, consequences, and remedies, *Phys. Rev.*, E 71, 066 110, 2005.
- Lee, Y. W.: *Statistical Theory of Communication*, John Wiley & Sons, New York, NY, 1960.
- Love, J. J.: Magnetic monitoring of Earth and space, *Physics Today*, 61, 31–37, 2008.
- Love, J. J.: Missing data and the accuracy of magnetic-observatory hour means, *Ann. Geophys.*, 27, 3601–3610, doi:10.5194/angeo-27-3601-2009, 2009.
- Macmillan, S. and Droujinina, A.: Long-term trends in geomagnetic daily activity, *Earth Planets Space*, 59, 391–395, 2007.
- Marsal, S. and Curto, J. J.: A new approach to the hourly mean computation problem when dealing with missing data, *Earth Planets Space*, 61, 945–956, 2009.
- Martini, D. and Mursula, K.: Correcting the geomagnetic IHV index of the Eskdalemuir observatory, *Ann. Geophys.*, 24, 3411–3419, doi:10.5194/angeo-24-3411-2006, 2006.
- Meteorological Office: *Hourly Values from Autographic Records: Geophysical Section. 1912*, His Majesty’s Stationary Office, London, UK, 1914.

- Meteorological Office: The Observatories' Year Book 1931, His Majesty's Stationary Office, London, UK, 1933.
- Meteorological Office: The Observatories' Year Book 1932, His Majesty's Stationary Office, London, UK, 1934.
- Olsen, N.: Natural sources for electromagnetic induction studies, in: *Encyclopedia of Geomagnetism and Paleomagnetism*, edited by: Gubbins, D. and Herrero-Bervera, E., pp. 743–746, Springer-Verlag, New York, NY, 2007.
- Parkinson, W. D.: *Introduction to Geomagnetism*, Scottish Academic Press, Edinburgh, UK, 1983.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: *Numerical Recipes*, Cambridge Univ. Press, Cambridge, UK, 1992.
- Prölss, G. W.: *Physics of the Earth's Space Environment*, Springer-Verlag, Berlin, Germany, 2004.
- Rasson, J. L.: INTERMAGNET, in: *Encyclopedia of Geomagnetism and Paleomagnetism*, edited by: Gubbins, D. and Herrero-Bervera, E., pp. 715–717, Springer-Verlag, New York, NY, 2007.
- Sabaka, T. J., Olsen, N., and Purucker, M. E.: Extending comprehensive models of the Earth's magnetic field with Ørsted and CHAMP data, *Geophys. J. Int.*, 159, 521–547, doi:10.1111/j.1365-246X.2004.02421.x, 2004.
- Schmidt, A.: *Ergebnisse der magnetischen Beobachtungen In Potsdam im Jahre 1905*, vol. 196, Veröff. K. Preuss. Meteor. Inst., Berlin, 1908.
- Svalgaard, L. and Cliver, E. W.: Interhourly variability index of geomagnetic activity and its use in deriving the long-term variation of solar wind speed, *J. Geophys. Res.*, 112, A10111, doi:10.1029/2007JA012437, 2007.